MULTIMEDIA UNIVERSITY

STUDENT ID NO

MULTIMEDIA UNIVERSITY

FINAL EXAMINATION

TRIMESTER 2, 2018/2019

TPR2251 – PATTERN RECOGNITION

(All sections / Groups)

01 MARCH 2019
03:00 p.m. – 05:00 p.m.
(2 Hours)

INSTRUCTIONS TO STUDENTS

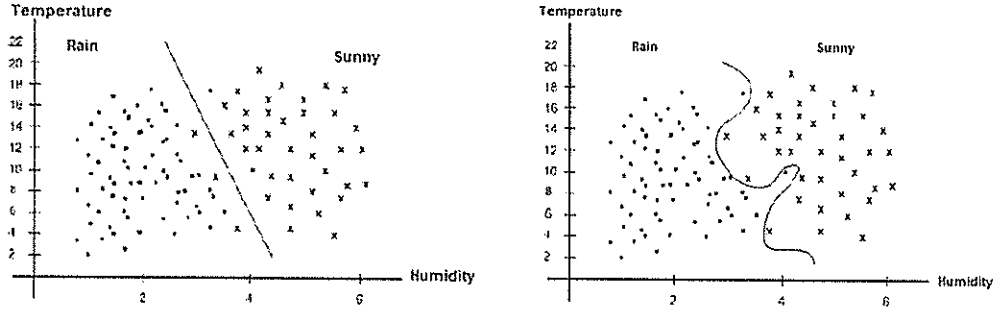
1. This question paper consists of 5 pages with 5 questions only.
2. Attempt **ALL** questions. All questions carry equal marks and the distribution of the marks for each question is given.
3. Please print all your answers in the answer booklet provided.

QUESTION 1

(a) Car plate recognition is a popular application of pattern recognition techniques. Explain the FOUR important processes in a car plate recognition system.

[4 marks]

(b) Suppose you have designed two classifiers shown in Figure 1 to predict whether it will rain the next day based on today's temperature and humidity measures. Which model (Model 1 or Model 2) will be more appropriate for the weather prediction problem? Justify your answer.



Model 1 Model 2

Figure 1. Two classifiers (Model 1 and 2) for weather prediction.

[2 marks]

(c) A genetic test has a probability 0.96 of giving a positive result when applied to a person suffering from a certain genetic defect, and a probability 0.08 of giving a false positive when applied to a non-sufferer. It is estimated that 0.2% of the population suffers from the genetic defect. Suppose that the test is now administered to a person. Calculate the following probabilities that:

(i) the test result will be positive. [1 mark]
(ii) given a positive result, the person suffers from the defect. [1 mark]
(iii) given a negative result, the person is a non-sufferer. [1 mark]
(iv) the person will be misclassified. [1 mark]

Continued .....

QUESTION 2

(a) Given the following two-class pattern recognition problem:

You are designing a classifier to separate two kinds of fish: tenggiri and cencaru. You need to predict what type of fish will emerge next on the conveyer belt and that the sequence of types of fish appears to be random.

(i) What are prior probabilities in the problem? [1 mark]
(ii) Name FOUR features that can be used in the classifier. [2 marks]

(b) In a two-class problem with a single feature x, the likelihood functions for ω_1 and ω_2 are given as follows:

$$p(x|\omega_1) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x-6)^2}{3}}$$
$$p(x|\omega_2) = \frac{1}{\sqrt{2\pi}} e^{-\frac{(x+3)^2}{2}}$$

If $P(\omega_1) = 2/3$ and $P(\omega_2) = 1/3$, find the minimum error classification and the corresponding expected classification error for $x = 2$. [3 marks]

(c) In a three-class, two-dimensional problem, the feature vectors in each class are normally distributed with a covariance matrix

$$\Sigma = \begin{bmatrix} 1.2 & 0.4 \\ 0.4 & 1.8 \end{bmatrix}$$

The mean vectors for each class are $[0.1 \quad 0.1]^T$, $[2.1 \quad 1.9]^T$, and $[-1.5 \quad 2.0]^T$, respectively. Assuming that the classes are equiprobable, classify the feature vector $[1.6 \quad 1.5]^T$ according to the Bayes minimum error probability classifier. [4 marks]

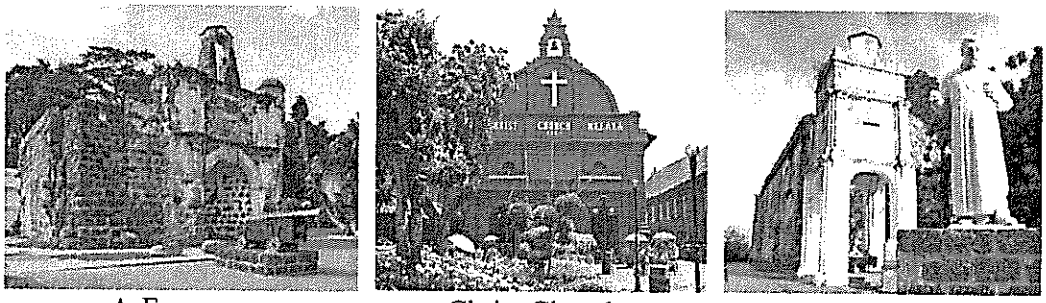
Continued .....

QUESTION 3

(a) What is dimension reduction? Explain why dimension reduction is crucial in a pattern recognition system like face recognition. [3 marks]

(b) In your opinion, how is *k*-NN different from *k*-means clustering? Explain your answer. [3 marks]

(c) Suppose you are building a classifier using Principal Component Analysis (PCA) to predict landmark buildings in Melaka. You are given some input training images as shown in Figure 2. What pre-processing steps are required on these images if you want to get a reasonably good performance from the PCA algorithm? List TWO steps. [2 marks]



A Famosa Christ Church St Paul Hill

Figure 2. Three landmark buildings in Melaka.

(d) Given a snapshot showing the scatter plot of two features, X1 and X2, comprising of two classes. Do you think Principal Component Analysis (PCA) can produce a good classification result on this data? Justify your answer.

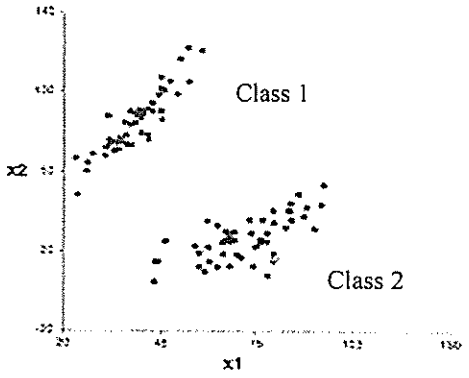


Figure 3. Scatter plot for two features, X1 and X2.

[2 marks]

Continued .....

QUESTION 4

(a) What is the difference between a generative and discriminative model? Explain. [2 marks]

(b) You want to cluster the following eight points into three clusters using k-means algorithm: $A_1(2,10)$, $A_2(2,5)$, $A_3(8,4)$, $A_4(5,8)$, $A_5(7,5)$, $A_6(6,4)$, $A_7(1,2)$, and $A_8(4,9)$. The distance function between two points, $a = (x_1, y_1)$ and $b = (x_2, y_2)$, is given by $d(a,b) = |x_2 - x_1| + |y_2 - y_1|$.

(i) Suppose the initial cluster centers are A_1, A_4, and A_7. Find the cluster centers after the first iteration. [4 marks]

(ii) The cluster centers do not change anymore after the forth iteration. Predict the class of a new data point $X(6,10)$ if the new cluster centers are $C_1(2,4)$, $C_2(3,6)$, and , $C_3(7,9)$, respectively. [2 marks]

(c) Given the following six points with their two corresponding attributes:

Points	Attribute 1	Attribute 2
x_1	0.4005	0.5306
x_2	0.2148	0.3854
x_3	0.3457	0.3156
x_4	0.2652	0.1875
x_5	0.0789	0.4139
x_6	0.4548	0.3022

The pair-wise distance matrix for the points are given by,

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	0.0000	0.2357	0.2218	0.3688	0.3421	0.2347
x_2	0.2357	0.0000	0.1483	0.2042	0.1388	0.2540
x_3	0.2218	0.1483	0.0000	0.1513	0.2843	0.1100
x_4	0.3688	0.2042	0.1513	0.0000	0.2932	0.2216
x_5	0.3421	0.1388	0.2843	0.2932	0.0000	0.3921
x_6	0.2347	0.2540	0.1100	0.2216	0.3921	0.0000

Plot the dendogram implemented using complete linkage proximity function. [2 marks]

Continued .....

QUESTION 5

(a) With a supervised learning algorithm, we can specify target output values but we may never get close to those targets at the end of learning. Give TWO reasons why this might happen. [2 marks]

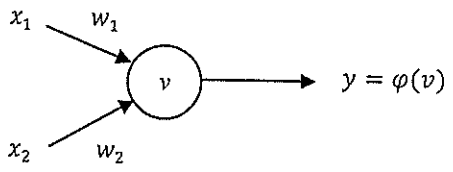
(b) Consider a SVM classifier with a polynomial kernel of degree two and answer the questions below.

(i) Suppose you applied this model to your data and found that it perfectly fits the data (e.g., both training and testing accuracies are 100%). What would you observe if the degree of the polynomial kernel is increased from two to five? [1 mark]

(ii) After increasing the complexity of the model, you found that the training accuracy was still 100%. What are the reasons behind this? Give TWO reasons. [2 marks]

(c) What is a training set and how is it used to train neural networks? [2 marks]

(d) Logical operators (i.e. NOT, AND, OR, and XOR) are the building blocks of any computational device. Logical functions return only two possible values, true or false, based on the truth or false values of their arguments. Verify that the AND logical function can be implemented using the following single-unit with inputs:



where $w_1 = 1$ and $w_2 = 1$ represent the weights, respectively, and v is the threshold level. The activation is given by,

$$\varphi(v) = \begin{cases} 1 & if\ v \geq 2 \\ 0 & otherwise \end{cases}$$

Hint: You may denote the truth value by 1 and false by 0.

[3 marks]

End of Page